
Developing optimal search strategies for detecting clinically sound treatment studies in EMBASE*

By Sharon S.-L. Wong, PhD
sharonw@ryerson.ca
Assistant Professor

School of Nutrition
Faculty of Community Services
Ryerson University
Toronto, Ontario M5B 2K3
Canada

Nancy L. Wilczynski, MSc
wilczyn@mcmaster.ca
Research Associate
Health Information Research Unit
Department of Clinical Epidemiology and Biostatistics

R. Brian Haynes, MD, PhD
bhaynes@mcmaster.ca
Professor of Clinical Epidemiology and Medicine
For the Hedges Team
Health Information Research Unit
Department of Clinical Epidemiology and Biostatistics and
Department of Medicine
Faculty of Health Sciences

McMaster University
Hamilton, Ontario, L8N 3Z5
Canada

Objective: The ability to accurately identify articles about therapy in large bibliographic databases such as EMBASE is important for researchers and clinicians. Our study aimed to develop optimal search strategies for detecting sound treatment studies in EMBASE in the year 2000.

Methods: Hand searches of journals were compared with retrievals from EMBASE for candidate search strategies. Six trained research assistants reviewed fifty-five journals indexed in EMBASE and rated articles using purpose and quality indicators. Candidate search strategies were developed for identifying treatment articles and then tested, and the retrievals were compared with the hand-search data. The operating characteristics of the strategies were calculated.

Results: Three thousand eight hundred fifty articles were original studies on treatment, of which 1,256 (32.6%) were methodologically sound. Combining search terms revealed a top performing strategy (random.tw. OR clinical trial.mp. OR exp health care quality) with sensitivity of 98.9% and specificity of 72.0%. Maximizing specificity, a top performing strategy (double-blind.mp. OR placebo.tw. OR blind:.tw.) achieved a value over 96.0%, but with compromised sensitivity at 51.7%. A 3-term strategy achieved the best optimization of sensitivity and specificity (random.tw. OR placebo.mp. OR double-blind.tw.), with both these values over 92.0%.

Conclusion: Search strategies can achieve high performance for retrieving sound treatment studies in EMBASE.

BACKGROUND

The ongoing tremendous growth in the medical literature, with an estimated publication rate of over two million new articles per year [1], ensures that the task of keeping up to date with the latest health care knowledge can be complex and labor intensive for researchers and impractical for the busy clinician [2, 3]. Nevertheless, determining the current state of knowledge is essential to both.

For treatment, studies of the randomized controlled trial design are deemed the most valid for clinical intervention studies [4]. Articles about therapy make up the largest proportion of clinically important articles in primary health care journals [5]. Furthermore, with the rising emphasis on evidence-based practice [4], the need is accentuated for simple, fast, reliable, and inexpensive ways of retrieving evidence both relevant to clinical practice and scientifically sound.

Increasingly, clinicians are performing their own searches and turning to large online biomedical literature databases to find information to support evidence-based decision making, such as choosing treatment options [6–8]. The success of a search is highly susceptible to the choice of database [9]. MEDLINE is often searched as the first-choice database because it provides free access to a broad range of biomedical literature coverage including nursing, dentistry, communication disorders, paramedical professions, population and reproductive biology, and clinical and experimental medicine. EMBASE searches, by comparison, are not free but complement MEDLINE in several ways, by providing greater coverage of European and non-English-language publications and broader coverage of key topics such as pharmaceuticals, psychiatry, toxicology, and alternative medicine [10]. The estimated overlap between MEDLINE and EMBASE is only 30% to 50% [11–14]. Searchers comparing the two databases have concluded that relevant information would inevitably be missed if only one of the databases were searched [6, 11, 15]. Those wanting the most comprehensive coverage, particularly of drugs and other therapeutic regimens, would likely benefit from searching EMBASE in conjunction with MEDLINE.

Sifting through the vast amount of information in large, general bibliographic databases to locate the best health care evidence can be taxing, especially when most articles in these databases are clinically irrelevant, are ambiguous, and have poor methodologic quality [5] and when clinicians feel uncertain about their search skills [16]. One solution for improving accuracy of retrieval of studies of various designs and contexts in online bibliographic databases is the use of “hedges” or filters, including search strategies consisting of indexing terms and textwords [17]. Although search strategies are not a completely satisfactory solution (for example, precision will remain generally

low in large, multipurpose databases containing a low concentration of relevant articles), they do help to narrow the search to a clinically important and sound subset of articles and have been commended to be influential and fundamental to information retrieval for evidence-based practice [17, 18].

Methodologic search strategies have been developed for improving the accuracy of retrieving treatment studies in MEDLINE [19–24]. Search strategies developed for searching MEDLINE cannot be directly translated for use in other databases such as EMBASE, because indexing practices vary and equivalent thesaurus terms do not necessarily exist across databases. Empirical studies on developing search strategies in EMBASE are scarce; the authors have identified only one study that developed search strategies for retrieving diagnostic studies in EMBASE [25]. To our knowledge, no empirical studies have been done to develop search strategies for detecting treatment studies in EMBASE.

In the 1990s, we developed MEDLINE search strategies on a small subset of 10 journals for articles pertaining to therapy, diagnosis, prognosis, and causation [26, 27]. We expanded and updated our work using data from 161 journals indexed in MEDLINE for the publishing year 2000 [28–33]. These strategies have been adapted for use in the Clinical Queries interface of MEDLINE <<http://www.ncbi.nlm.nih.gov/entrez/query/static/clinical.html>> and in other searching arenas [34].

Furthering our earlier work, we report here the retrieval properties of single and combined terms for identifying methodologically sound studies on the prevention and treatment of health disorders in EMBASE. Although we acknowledge that indexing and terminology can evolve over time, we confined our manual search to the year 2000, having previously established the robustness of search strategies across publication periods (1991 and 2000) in MEDLINE [35].

METHODS

We compared the operating characteristics of methodologic search strategies in EMBASE with a manual review of each article in each issue of fifty-five journals (Table 1) for the publishing year 2000. To evaluate EMBASE strategies designed to retrieve treatment studies, index terms and textwords related to research design features were run as search strategies. We treated the search strategies as “diagnostic tests” for sound treatment studies and the manual review of the literature as the “gold standard.” The sensitivity (or “recall”), specificity, precision, and accuracy of EMBASE searches were determined as shown in Figure 1. For example, for each EMBASE search strategy designed to retrieve sound treatment studies, sensitivity was defined as the proportion of relevant, high-quality articles retrieved; specificity as the proportion of low-quality articles not retrieved; precision as the proportion of relevant and high-quality articles retrieved; and accuracy as the proportion of all correctly classified articles.

* Source of funding: National Library of Medicine, USA (grant no. 5R01LM06866-04).

Table 1

List of fifty-five journals indexed in EMBASE manually reviewed in the year 2000

<i>Acta Obstetrica et Gynecologica Scandinavica</i>	<i>Diabetes Care</i>
<i>American Journal of Cardiology</i>	<i>Diabetic Medicine</i>
<i>American Journal of Epidemiology</i>	<i>Fertility & Sterility</i>
<i>American Journal of Gastroenterology</i>	<i>Heart</i>
<i>American Journal of Medicine</i>	<i>International Journal of Geriatric Psychiatry</i>
<i>American Journal of Obstetrics & Gynecology</i>	<i>JAMA</i>
<i>American Journal of Psychiatry</i>	<i>Journal of Clinical Oncology</i>
<i>American Journal of Respiratory & Critical Care Medicine</i>	<i>Journal of Clinical Psychiatry</i>
<i>Annals of Emergency Medicine</i>	<i>Journal of Clinical Psychopharmacology</i>
<i>Annals of Internal Medicine</i>	<i>Journal of Consulting & Clinical Psychology</i>
<i>Annals of Surgery</i>	<i>Journal of Family Practice</i>
<i>Archives of Disease in Childhood</i>	<i>Journal of Infectious Diseases</i>
<i>Archives of Disease in Childhood Fetal & Neonatal Edition</i>	<i>Journal of Pediatrics</i>
<i>Archives of Family Medicine</i>	<i>Journal of Rheumatology</i>
<i>Archives of General Psychiatry</i>	<i>Journal of the American College of Cardiology</i>
<i>Archives of Internal Medicine</i>	<i>Journal of the American Geriatrics Society</i>
<i>Archives of Pediatrics & Adolescent Medicine</i>	<i>Lancet</i>
<i>Archives of Physical Medicine & Rehabilitation</i>	<i>Neurology</i>
<i>BJOG</i>	<i>New England Journal of Medicine</i>
<i>BMJ</i>	<i>Obstetrics & Gynecology</i>
<i>British Journal of General Practice</i>	<i>Pain</i>
<i>British Journal of Psychiatry</i>	<i>Pediatrics</i>
<i>British Journal of Surgery</i>	<i>Radiology</i>
<i>Cancer</i>	<i>Social Science & Medicine</i>
<i>Cancer Nursing</i>	<i>Spine</i>
<i>Chest</i>	<i>Stroke</i>
<i>Circulation</i>	<i>Thorax</i>
<i>Critical Care Medicine</i>	

Six research assistants assessed all articles for each issue of 55 journals for the year 2000. For articles in 7 purpose categories (causation, prognosis, diagnosis, treatment, economics, clinical prediction, and reviews), methodologic criteria were applied to determine if the article was scientifically sound. We used purpose category definitions to classify qualitative and cost studies but did not apply methodologic criteria to these types of studies. Purpose category definitions and corresponding methodologic rigor criteria have previously been published [36]. Original articles (of interest to the health care of humans) pertaining to the prevention and treatment of diseases and health disorders were required to meet these methodologic criteria: random allocation of participants to comparison groups, outcome assessment of at least 80% of participants entering the investigation in 1 major analysis at

any given follow-up assessment, and analysis consistent with study design. Research staff underwent training and intensive calibration before reviewing the 2000 literature, and interrater reliability (analyzed by the *kappa* statistic) for application of all methodologic criteria exceeded 80% beyond chance for all purpose categories [36].

The journals were selected using an iterative process based on recommendations of clinicians and librarians, Science Citation Index impact factors provided by the Institute for Scientific Information, and their ongoing yield of studies and reviews of scientific merit and clinical relevance for the disciplines of internal medicine, general medical practice, mental health, and general nursing practice (full list of journals available upon request from authors). Search strategies were developed on a 55-journal subset (Table 1), selected from 135 EMBASE clinical journals initially reviewed because of their ongoing yield of sound and clinically relevant articles. We had previously developed search strategies in MEDLINE using 161 journals indexed in MEDLINE [28–33] but found that the search strategies were robust in smaller journal subsets and that computation time was substantially decreased. We also found that, when strategies were developed in 60% of the database and validated in the remaining 40%, there were no statistical differences in performance. Thus, we developed search strategies for EMBASE using all data from the 55 journals.

To construct a comprehensive set of search terms, we compiled an initial list of index terms and text-words and then sought input from clinicians and librarians in the United States and Canada through interviews with known searchers and requests at meetings and conferences. Individuals were asked to iden-

Figure 1

Formula for calculating the sensitivity, specificity, precision, and accuracy of EMBASE searches for detecting clinically sound treatment studies.

		Manual review ("gold standard")	
		Meets criteria	Does not meet criteria
Search strategies ("diagnostic tests")	Detected	a	b
	Not detected	c	d
		a + c	b + d

Sensitivity = $a/(a+c)$; specificity = $d/(b+d)$; precision = $a/(a+b)$; accuracy = $(a+d)/(a+b+c+d)$. All articles classified during the manual review of the literature = $(a+b+c+d)$.

Table 2

Single terms with the best sensitivity (keeping specificity $\geq 50\%$), best specificity (keeping sensitivity $\geq 50\%$), and best optimization of sensitivity and specificity (based on the lowest possible absolute difference between sensitivity and specificity) for detecting treatment studies in EMBASE in 2000

Search term OVID search*	Sensitivity (95% CI) (number of high-quality treatment articles = 1,256)	Specificity (95% CI) (number of low-quality articles = 26,513)	Precision (95% CI) (denominator varies by search term)	Accuracy (95% CI) (total number of articles = 27,769)
Best sensitivity: random:.mp.	95.1% (94.0 to 96.3)	92.5% (92.2 to 92.8)	37.6% (35.9 to 39.3)	92.6% (92.3 to 92.9)
Best specificity: randomized.tw.	63.2% (60.6 to 65.9)	96.7% (96.5 to 96.9)	47.5% (45.1 to 49.9)	95.2% (94.9 to 95.4)
Best optimization of sensitivity and specificity: clinical trial:.mp.	88.3% (86.5 to 90.1)	88.0% (87.6 to 88.4)	25.9% (24.5 to 27.2)	88.0% (87.6 to 88.4)

* Search strategies are reported using Ovid's search engine syntax for EMBASE. : = truncation; mp = multiple posting (term appears in title, abstract, or subject heading); tw = textword (word or phrase appears in title or abstract).

tify which terms or phrases they used when searching for qualitative studies of treatment, causation, prognosis, diagnosis, economics, clinical prediction guides, reviews, and costs. We compiled a list of 5,385 terms, of which 4,843 were unique and 3,524 returned results (list of tested terms available on request from authors). For example, tested search terms included the textwords, "random," "randomized trial," "efficacy," and "intention to treat"; the index term, "clinical trial"; and the index term, "treatment outcome," exploded (exploded search terms also retrieve records indexed using more specific, narrower terms of the entered term).

Individual search terms with sensitivity greater than 25% and specificity greater than 75% for a given purpose category were incorporated into the development of search strategies that included a combination of 2 or more terms. All combinations of terms used the Boolean "OR," for example, "random OR controlled." The Boolean "AND" was not used because this strategy invariably compromised sensitivity, because it limited the search by reducing the number of citations retrieved. For the development of multiple-term search strategies to either optimize sensitivity or specificity, we tested all 2-term search strategies with sensitivity of at least 75% and specificity of at least 50%. For optimizing accuracy, 2-term search strategies with accuracy greater than 75% were considered for multiple-term development. Seven thousand one hundred sixty-four search strategies were tested in the development of treatment search filters.

Logistic regression approaches to developing search strategies were used when deriving treatment and prognostic hedges for MEDLINE. These logistic regression approaches, compared with the Boolean approach described above, did not improve search strategy performance. Hence, for the other purpose categories and databases including EMBASE, only the Boolean approach was used for search strategy development.

RESULTS

Indexing information was downloaded from EMBASE for 27,769 articles from the 55 journals from 2000 that were hand-searched. Of these, 3,850 were classified as

original articles of treatment, of which 1,256 (32.6%) were methodologically sound. Search strategies were developed using all 27,769 articles. Thus, the strategies were tested for their ability to retrieve high-quality treatment articles from all other articles, including both poor-quality treatment studies and all nontreatment studies.

Table 2 shows the single terms with the best sensitivity, best specificity, and best optimization of sensitivity and specificity for detecting sound treatment studies in EMBASE in 2000. The term, "random:.mp.," achieved the best sensitivity at 95.1%; even with sensitivity maximized, specificity was high at 92.5%. The single term, "randomized.tw.," achieved the best specificity at 96.7%, albeit with a clear but expected reduction in sensitivity (63.2%). With specificity maximized, precision was increased to just over 47% (an absolute increase in precision of about 10% compared with the best sensitivity single term). The single term, "clinical trial:.mp.," achieved the best balance of sensitivity (88.3%) and specificity (88.0%).

The operating characteristics of top performing combination strategies are shown in Table 3. The 3-term strategy, "random.tw. OR clinical trial:.mp. OR exp health care quality," yielded the best sensitivity (almost 99%) and had a specificity of 72.0%. Compared with the single term with the best sensitivity, "random:.mp." (95.1% sensitivity, 92.5% specificity, and 37.6% precision), the 3-term strategy with the best sensitivity achieved an absolute gain in sensitivity of only 3.8% but with quite substantive absolute losses in specificity (20.5%) and precision (23.3%).

For the 3-term strategy with the best sensitivity, replacing the term, "exp health care quality," with the term, "exp treatment outcome," led to an absolute increase in specificity of 5.8% (72.0% to 77.8%), with only a small 0.2% decrease in sensitivity (98.9% to 98.7%) (Table 3).

The 3-term strategy, "double-blind:.mp. OR placebo:.tw. OR blind:.tw.," yielded the best specificity at 96.7%, but with a definite trade-off in sensitivity, which lowered to 51.7%. Yet with specificity maximized, a relatively remarkable rise was seen in precision, which reached 42.8%. Compared with the 3-term strategy with the best sensitivity, this represented an absolute increase in precision of 28.5%. When search

Table 3

Combination of terms with the best sensitivity (keeping specificity $\geq 50\%$), best specificity (keeping sensitivity $\geq 50\%$), and best optimization of sensitivity and specificity (based on absolute [sensitivity-specificity] $< 1\%$) for detecting treatment studies in EMBASE in 2000

Search strategy OVID search*	Sensitivity (95% CI) (number of high-quality treatment articles = 1,256)	Specificity (95% CI) (number of low-quality articles = 26,513)	Precision (95% CI) (denominator varies by search term)	Accuracy (95% CI) (total number of articles = 27,769)
Best sensitivity: random:.tw. OR clinical trial:.mp. OR exp health care quality	98.9% (98.3 to 99.5)	72.0% (71.4 to 72.5)	14.3% (13.6 to 15.1)	73.2% (72.7 to 73.7)
Small drop in sensitivity with a substantive gain in specificity: random:.tw. OR clinical trial:.mp. OR exp treatment outcome	98.7% (98.1 to 99.4)	77.8% (77.3 to 78.3)	17.4% (16.5 to 18.3)	78.7% (78.3 to 79.2)
Best specificity: double-blind:.mp. OR placebo:.tw. OR blind:.tw.	51.7% (48.9 to 54.4)	96.7% (96.5 to 96.9)	42.8% (40.3 to 45.3)	94.7% (94.4 to 95.0)
Best optimization of sensitivity and specificity: random:.tw. OR placebo:.mp. OR double-blind:.tw.	94.5% (93.3 to 95.8)	92.6% (92.3 to 92.9)	37.8% (36.1 to 39.5)	92.7% (92.4 to 93.0)

* Search strategies are reported using Ovid's search engine syntax for EMBASE. : = truncation; tw = textword (word or phrase appears in title or abstract); mp = multiple posting (term appears in title, abstract, or subject heading); exp = exploded subject heading.

terms were combined to optimize sensitivity and specificity, these values exceeded 92%. Accuracy was driven by specificity and, for all top-performing terms, these percentages were very similar.

DISCUSSION

We have developed search strategies that can assist clinicians and researchers in retrieving methodologically sound treatment studies. Choice of search strategy should be made by weighing the most appropriate trade-off between sensitivity and specificity to best fulfill the purpose of the search. For example, if the purpose of the search is to identify as complete a set as possible of relevant randomized controlled trials for a systematic review, a high sensitivity strategy would be most suitable. For maximum coverage, however, even more widespread albeit more labor-intensive approaches might be considered. For example, searching more than one database such as EMBASE plus MEDLINE, hand-searching bibliographies of relevant articles and key journals, and contacting experts and pharmaceutical companies could be done. On the other hand, if the purpose of the search is to efficiently retrieve several key articles, a high specificity strategy that reduces the number of nuisance hits might be suitable.

Our 3-term results in Table 3 show that, by using the best specificity instead of the best sensitivity strategy, sensitivity drops 47.2% from 98.9% (with most sensitive strategy) to 51.7% (with most specific strat-

egy), meaning that almost 1 of every 2 clinically relevant and sound articles would be missed. Bachmann and colleagues [18] have raised uncertainty that the articles retrieved in searches are necessarily a random selection of those available. Therefore, using the most specific strategy (even though it has greater precision), which lacks good sensitivity, can be risky because a biased representation of the knowledge based on only half of the available evidence cannot be excluded. The strategy that best optimizes sensitivity and specificity has better precision than the most sensitive strategy, with a much smaller compromise in sensitivity than the best specificity strategy. For searches of broad content areas with anticipated high citation yields, but requiring completeness, the best optimization strategy might be the most sensible option.

For example, suppose a search is done in EMBASE to identify articles on the effectiveness of therapy with herbal medicine. The search might begin with the single content term, "herbal medicine," which would give an enormous yield of 5,696 articles (Table 4). Combining this initial search with the combination strategy with the *best sensitivity* that we developed reduces the yield more than five-fold to 1,115 articles, but this is still a cumbersome subset. Nevertheless, the 1,115 retrievals might be worth sifting through for the purpose of being as inclusive as possible. Alternatively, combining the initial search with the combination strategy with the *best specificity* that we developed reduces the yield to a far more manageable 238 articles (one-fifth the yield of the most sensitive strategy), although the

Table 4

Example: EMBASE (1980 to 2004, week 48) searches to identify articles on the therapeutic effectiveness of herbal medicine

Search strategy	Terms	Results
Content term alone	Herbal medicine	5,696 articles
Content term AND <i>best sensitivity</i> combination strategy	Herbal medicine AND "random:.tw. OR clinical trial:.mp. OR exp health care quality"	1,115 articles
Content term AND <i>best specificity</i> combination strategy	Herbal medicine AND "double-blind:.mp. OR placebo:.tw. OR blind:.tw."	238 articles
Content term AND <i>best optimization of sensitivity and specificity</i> combination strategy	Herbal medicine AND "random:.tw. OR placebo:.mp. OR double-blind:.tw."	427 articles

possibility exists that some important articles are not detected. Finally, choosing instead to combine the initial search with the combination strategy with the *best optimization of sensitivity and specificity* yields an intermediate number of 427 hits. Note that a quick scan of the 238 articles detected by the *best specificity* strategy reveals at least several articles on the effectiveness of herbal therapies—such as ginkgo, St. John's wort, ginseng, echinacea, saw palmetto, kava, and Chinese herbal medicine—that have been included in *ACP Journal Club* [37–39]. Because *ACP Journal Club* includes only articles that meet basic criteria for clinical relevance and scientific merit, this search successfully identified at least several relevant and sound articles.

Although we developed our filters to retrieve treatment studies meeting certain criteria for methodologic rigor, it does not necessarily imply that the retrieved articles will have been done using the best methodologic standards. Ultimately, the end user has the responsibility for appraising the retrieved literature for quality and relevance before applying it to clinical practice. The quality of randomized controlled trials accepted for publication, for example, differs by journal [40], which may be attributable to variations in the rigor of the peer-review process and instructions given to authors and reviewers. Journal editors, by giving explicit instructions to authors to fully report their methodology, can help improve the accuracy of the indexing process in bibliographic databases and facilitate the reader's appraisal of study quality.

Dissimilarities in bibliographic databases (e.g., coverage of different sets of journals, variability in indexing practices and thesaurus terms, and use of different search engines) make direct comparisons of our top-performing treatment filters for EMBASE with those for MEDLINE not meaningful. For example, one of the terms in the most sensitive combination strategy, "health care quality," is a subject heading in EMBASE that is not supported as an index term in MEDLINE. Therefore, direct testing of this EMBASE strategy in the MEDLINE database is not possible. Conceivably, optimal filters in one database may not be top-performing filters in another database; hence, filters need to be developed for specific use in the intended database.

Search strategies are a helpful but imperfect solution to accurate literature retrieval in large online bibliographic databases. Even our top-performing strategies have generally low precision because EMBASE is such a large and broad-ranging database. Because precision in large multipurpose databases is inevitably low due to the small concentration of relevant articles, it is especially important that terminology (e.g., methodologic terminology) used in therapeutic studies be as accurate, explicit, and consistent as possible to facilitate the indexing process and improve the success of a search. Variations in the meticulousness of indexing quality exist in bibliographic databases [18]. Other factors that can also impact the success of a search are the formation of a well-defined clinical question and

its translation into a searchable strategy and the skill of the searcher.

Our study shows that the retrieval of methodologically sound articles on therapy and prevention in EMBASE can be enhanced by the use of several search filters. Although beneficial filters exist, such as those reported here, further work is needed to improve dissemination and publication of filters so that clinicians, researchers, and librarians are not only aware of them, but also have greater knowledge and proficiency in using them effectively.

CONCLUSION

Several search strategies can achieve high performance in retrieving methodologically sound studies on the prevention and treatment of diseases and health disorders in EMBASE. The optimal trade-off between sensitivity and specificity should be assessed based on the purpose of the search.

ACKNOWLEDGMENTS

The Hedges Team includes Angela Eady, R. Brian Haynes, Susan Marks, Kathleen Ann McKibbin, Doug Morgan, Cindy Walker-Dilks, Stephen Walter, Stephen Werre, Nancy Wilczynski, and Sharon Wong, all in the Health Information Research Unit, Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Ontario, Canada.

REFERENCES

1. LEE KP, SCHOTLAND M, BACCHETTI P, BERO LA. Association of journal quality indicators with methodological quality of clinical research articles. *JAMA* 2002 Jun 5;287(21):2805–8.
2. BIRKEN CS, PARKIN PC. In which journals will pediatricians find the best evidence for clinical practice? *Pediatrics* 1999 May;103(5 pt.1):941–7.
3. ALPER BS, HAND JA, ELLIOTT SG, KINKADE S, HAUN MJ, ONION DK, SKLAR BM. How much effort is needed to keep up with the literature relevant for primary care? *J Med Libr Assoc* 2004 Oct;92(4):429–37.
4. SACKETT DL, STRAUS SE, RICHARDSON WS, ROSENBERG W, HAYNES RB. Evidence-based medicine. how to practice and teach EBM. 2nd ed. London, UK: Churchill Livingstone, 2000.
5. MCKIBBIN KA, WILCZYNSKI NL, HAYNES RB. What do evidence-based secondary journals tell us about the publication of clinically important articles in primary healthcare journals? *BMC Med* 2004 Sep 6;2:33.
6. BRETTLE AJ, LONG AF. Comparison of bibliographic databases for information on the rehabilitation of people with severe mental illness. *Bull Med Libr Assoc* 2001 Oct;89(4):353–62.
7. FANGERAU H. Finding European bioethical literature: an evaluation of the leading abstracting and indexing services. *J Med Ethics* 2004 Jun;30(3):299–303.
8. WESTBROOK JI, GOSLING AS, COIERA E. Do clinicians use online evidence to support patient care? a study of 55,000 clinicians. *J Am Med Inform Assoc* 2004 Mar–Apr;11(2):113–20. Epub 2003 Dec 07.
9. McDONALD S, TAYLOR L, ADAMS C. Searching the right

- database. a comparison of four databases for psychiatry journals. *Health Libr Rev* 1999 Sep;16(3):151-6.
10. EMBASE. STN database summary sheet. [Web document]. EMBASE. [cited 16 Nov 2004]. <<http://www.cas.org/ONLINE/DBSS/embase.html>>.
 11. TOPFER LA, PARADA A, MENON D, NOORANI H, PERRAS C, SERRA-PRAT M. Comparison of literature searches on quality and costs for health technology assessment using the MEDLINE and EMBASE databases. *Int J Technol Assess Health Care* 1999 Spring;15(2):297-303.
 12. SMITH BJ, DARZINS PJ, QUINN M, HELLER RF. Modern methods of searching the medical literature. *Med J Aust* 1992 Nov 2;157(9):603-11.
 13. MINOZZI S, PISTOTTI V, FORNI M. Searching for rehabilitation articles on MEDLINE and EMBASE. an example with cross-over design. *Arch Phys Med Rehabil* 2000 Jun;81(6):720-2.
 14. SUAREZ-ALMAZOR ME, BELSECK E, HOMIK J, DORGAN M, RAMOS-REMUS C. Identifying clinical trials in the medical literature with electronic databases: MEDLINE alone is not enough. *Control Clin Trials* 2000 Oct;21(5):476-87.
 15. WOODS D, TREWHEELAR K. Medline and Embase complement each other in literature searches [letter]. *BMJ* 1998 Apr 11;316(7138):1166-7.
 16. ELY JW, OSHEROFF JA, EBELL MH, CHAMBLISS ML, VINSON DC, STEVERMER JJ, PIFER EA. Obstacles to answering doctors' questions about patient care with evidence: qualitative study. *BMJ* 2002 Mar 23;324(7339):710.
 17. JENKINS M. Evaluation of methodological search filters: a review. *Health Info Libr J* 2004 Sep;21(3):148-63.
 18. BACHMANN LM, CORAY R, ESTERMANN P, TER RIET G. Identifying diagnostic studies in MEDLINE: reducing the number needed to read. *J Am Med Inform Assoc* 2002 Nov-Dec;9(6):653-8.
 19. JADAD AR, MCQUAY HJ. A high-yield strategy to identify randomized controlled trials for systematic reviews. *Online J Curr Clin Trials* 1993 Feb 27;doc no.33.
 20. MARSON AG, CHADWICK DW. How easy are randomized controlled trials in epilepsy to find on Medline? the sensitivity and precision of two Medline searches. *Epilepsia* 1996 Apr;37(4):377-80.
 21. ADAMS CE, POWER A, FREDERICK K, LEFEBVRE C. An investigation of the adequacy of MEDLINE searches for randomized controlled trials (RCTs) of the effects of mental health care. *Psychol Med* 1994 Aug;24(3):741-8.
 22. DUMBRIGUE HB, ESQUIVEL JF, JONES JS. Assessment of MEDLINE search strategies for randomized controlled trials in prosthodontics. *J Prosthodont* 2000 Mar;9(1):8-13.
 23. NWOSU CR, KHAN KS, CHIEN PF. A two-term MEDLINE search strategy for identifying randomized trials in obstetrics and gynecology. *Obstet Gynecol* 1998 Apr;91(4):618-22.
 24. ROBINSON KA, DICKERSIN K. Development of a highly sensitive search strategy for the retrieval of reports of controlled trials using PubMed. *Int J Epidemiol* 2002 Feb;31(1):150-3.
 25. BACHMANN LM, ESTERMANN P, KRONENBERG C, TER RIET G. Identifying diagnostic accuracy studies in EMBASE. *J Med Libr Assoc* 2003 Jul;91(3):341-6.
 26. WILCZYNSKI NL, WALKER CJ, MCKIBBON KA, HAYNES RB. Assessment of methodologic search filters in MEDLINE. *Proc Annu Symp Comput Appl Med Care* 1993:601-5.
 27. HAYNES RB, WILCZYNSKI N, MCKIBBON KA, WALKER CJ, SINCLAIR JC. Developing optimal search strategies for detecting clinically sound studies in MEDLINE. *J Am Med Inform Assoc* 1994 Nov-Dec;1(6):447-58.
 28. WILCZYNSKI NL, HAYNES RB, HEDGES TEAM. Developing optimal search strategies for detecting clinically sound causation studies in MEDLINE. *AMIA Annu Symp Proc* 2003:719-23.
 29. WONG SS, WILCZYNSKI NL, HAYNES RB, RAMKISSOON-SINGH R, HEDGES TEAM. Developing optimal search strategies for detecting sound clinical prediction studies in MEDLINE. *AMIA Annu Symp Proc* 2003:728-32.
 30. HAYNES RB, WILCZYNSKI NL. Optimal search strategies for retrieving scientifically strong studies of diagnosis from MEDLINE: analytical survey. *BMJ* 2004 May 1;328(7447):1040. Epub 2004 Apr 08.
 31. WILCZYNSKI NL, HAYNES RB, HEDGES TEAM. Developing optimal search strategies for detecting clinically sound prognostic studies in MEDLINE: an analytic survey. *BMC Med* 2004 Jun 9;2:23.
 32. WONG SS, WILCZYNSKI NL, HAYNES RB, HEDGES TEAM. Developing optimal search strategies for detecting clinically relevant qualitative studies in MEDLINE. *Medinfo* 2004; 11(pt.1):311-6.
 33. MONTORI VM, WILCZYNSKI NL, MORGAN D, HAYNES RB, HEDGES TEAM. Optimal search strategies for retrieving systematic reviews from MEDLINE: analytical survey. *BMJ* 2005 Jan 8;330(7482):68-73.
 34. PUBLIC HEALTH RESOURCE UNIT, NATIONAL HEALTH SERVICE. Search filters. [Web document]. UK: The Service. [cited 16 Nov 2004]. <<http://www.phru.nhs.uk/casp/filters.htm>>.
 35. WILCZYNSKI NL, HAYNES RB, HEDGES TEAM. Robustness of empirical search strategies for clinical content in MEDLINE. *Proc AMIA Symp* 2002:904-8.
 36. WILCZYNSKI NL, MCKIBBON KA, HAYNES RB. Enhancing retrieval of best evidence for health care from bibliographic databases: calibration of the hand search of the literature. *Medinfo* 2001;10(pt.1):390-3.
 37. ERNST E. The risk-benefit profile of commonly used herbal therapies: ginkgo, St. John's wort, ginseng, echinacea, saw palmetto, and kava. *Ann Intern Med* 2002 Jan 1;136(1):42-53.
 38. WILT TJ, ISHANI A, STARK G, MACDONALD R, LAU J, MULROW C. Saw palmetto extracts for treatment of benign prostatic hyperplasia: a systematic review. *JAMA* 1998 Nov 11;280(18):1604-9.
 39. BENSOUSSAN A, TALLEY NJ, HING M, MENZIES R, GUO A, NGU M. Treatment of irritable bowel syndrome with Chinese herbal medicine: a randomized controlled trial. *JAMA* 1998 Nov 11;280(18):1585-9.
 40. BENDER JS, HALPERN SH, THANGAROOPAN M, JADAD AR, OHLSSON A. Quality and retrieval of obstetrical anaesthesia randomized controlled trials. *Can J Anaesth* 1997 Jan; 44(1):14-8.

Received December 2004; accepted July 2005